FRAMEWORK FOR DIAGNOSTIC ASSESSMENT OF MATHEMATICS

Edited by
Benő Csapó • Mária Szendrei

NEMZETI TANKÖNYVKIADÓ

Framework for Diagnostic Assessment of Mathematics

FRAMEWORK FOR DIAGNOSTIC ASSESSMENT OF MATHEMATICS

Edited by

Benő Csapó
Institute of Education, University of Szeged

Mária Szendrei
Department of Algebra and Number Theory, University of Szeged

Nemzeti Tankönyvkiadó
Budapest

Developing Diagnostic Assessment
Projekt ID: TÁMOP 3.1.9-08/1-2009-0001

National Development Agency
www.ujszechenyiterv.gov.hu
06 40 638 638

HUNGARY'S RENEWAL

The project is supported by the European Union
and co-financed by the European Social Fund.

Authors:
Benő Csapó, Csaba Csíkos, Katalin Gábri, Józsefné Lajos, Ágnes Makara,
Terezinha Nunes, Julianna Szendrei,
Mária Szendrei, Judit Szitányi, Lieven Verschaffel, Erzsébet Zsinkó

The chapters were reviewed by
József Kosztolányi and Ödön Vancsó

ISBN 978-963-19-7217-7

Nemzeti Tankönyvkiadó Zrt.
a Sanoma company

www.ntk.hu • Customer service: info@ntk.hu • Telephone: 06-80-200-788

Responsible for publication: János Tamás Kiss chief executive officer
Storing number: 42684 • Technical director: Etelka Vasvári Babicsné
Responsible editor: Katalin Fried • Technical editor: Tamás Kiss
Size: 28,6 (A/5) sheets • First edition, 2011

2

Mathematical Literacy and the Application of Mathematical Knowledge

Csaba Csíkos
Institute of Education, University of Szeged

Lieven Verschaffel
Institute of Education, Katholieke Universiteit, Leuven

One of the most important sources of objectives for learning mathematics can be summarized as the needs coming from the society in general and from other disciplines, especially sciences. Therefore, mathematics as a discipline and as a school subject may shape students' minds in a way that they develop a disposition to use their mathematical knowledge in several different contexts including other school subjects and everyday out-of-school problem contexts.

The idea of describing how the mathematical knowledge achieved in schools can be applied in various contexts and problem spaces is at least of the same age as the emerging mathematical ideas. Therefore the general theoretical fundamentals of the application phenomena will be first shortly presented in this chapter. In the last centuries, in most European school systems mathematics as a school subject earned the position of having a central role in curricula. Since the Ratio Studiorum, when Christopher Clavius exerted his influence on making mathematics a standard part of the Jesuit core curriculum (see Smolarski, 2002), till today's core curricula in Europe, there is a continuous search for better ways in teaching and learning mathematics. The second part of this chapter will focus on some assessment considerations about the applications of mathematics.

In the third part of this chapter, the characteristics and role of classroom mathematics tasks will be analyzed with a special emphasis on word problems. It is the classroom practice and culture that shape students' beliefs about and approaches of different types of word problems. Finally, we aim to provide a categorization of mathematical word problems in view of developing a diagnostic evaluation system of mathematical literacy.

Theoretical Considerations

In the history of mathematics and mathematics teaching there were continuous attempts and efforts made in order to bring evidence about the importance of mathematics in everyday life and in other sciences. These efforts have often been hindered by the dual nature of mathematics, i.e., the way mathematical results were published and communicated, and the way mathematical thinking and explorations have been actually performed.

The Nature of Mathematical Thinking

Mathematics is often associated with creating theorems, proofs and definitions. From ancient times, mathematical publications followed strict rules in presenting mathematical results. These rules are essentially the rules of deductive implications. The structure of many mathematical publications even today follows the sequence of definition – theorem – proof. However, as early as in the seventeenth century Descartes claimed that the ancient Greeks in fact yielded their theorems in an inductive way while they published their results according to strict deductive rules. The duality of how theorems are presented and how they have been achieved can even confuse laymen who often consider mathematician as people who create theorem and prove them. Nevertheless, Rickart (1996) emphasizes - following in Poincaré's and Hadamard's footsteps – that creativity plays an essential role in mathematical discovery. Conscious hard work and creative experiences go in tandem when doing mathematics. Although different facets of mathematical thinking go in tandem in doing mathematics, one or the other may noticeably appear, depending on the task to be solved. "Even inside the profession we classify ourselves as either theorists or problem solvers." (Guy, 1981, p. vii.) Ernest (1999) suggests keeping a balance between explicit propositional and tacit mathematical knowledge in educational contexts.

The key for understanding how school mathematics reflects different philosophical approaches can be found in Freudenthal's oeuvre. What students should learn in schools is to do mathematics and not primarily to accept the products of (mathematicians') mathematical activity. Doing mathematics requires students to gather experiences, form hypotheses, and above all, to learn to think mathematically. "The learner should reinvent mathematising rather than mathematics; abstracting rather than abstractions; schematizing rather than schemes; formalizing rather than formulae; algorithmising rather than algorithms; verbalizing rather than language… (Freudenthal, 1991, p. 49). Contrary to the historically developed DTP order (definition – theorem – proof), for mathematics lessons a reversed order should be applied: exploration, explanation, formalization (Hodgson, & Morandi, 1996).

A Mathematical Modeling Perspective

"The emergence of the discipline Mathematics Education in the beginning of the 20th century had a clear political motivation" (Sriraman & Törner, 2008, p. 668.) The main supporters of different movements were of economic nature. There are two mathematics education movements in the twentieth century that have strong influence on the principles and practices of even today's mathematics education. The New Math movement aimed at emphasizing mathematical structure through abstract concepts. Following the works of the Bourbaki group, the New Math movement has resulted in highly formalized textbooks, and initiated school curriculum and teacher education reforms. The New Math movement emphasized the *why*s and the deeper structure of mathematics, instead of mindless rigidity of traditional mathematics (Sriraman & Törner, 2008). That is why it is worth evaluating that movement in a more positive way instead of merely criticizing it from a postmodern math education perspective. This movement initiated studying the similarities between mathematical and psychological (hypothetico-deductive) structures as well.

The Realistic Mathematics Education (RME) movement is "a reaction to both the American New Math movement … and the then prevailing Dutch

… 'mechanistic mathematics education." (van den Heuvel-Panhuizen, 2001, p. 1). The RME grew out of Hans Freudenthal's initiations: founding the Wiskobas project (in Dutch: 'mathematics in primary school') and later the Freudenthal Institute, and at the same time fertilizing mathematics education with ideas such as that student should develop and apply concepts and tools for daily life problem situations that are meaningful for them (van den Heuvel-Panhuzien, 2003). As already indicated in the above-mentioned quotation from Freudenthal, realistic mathematics educations aims at the construction by children of their own mathematical knowledge, emphasizing human activity as mathematizing both within the mathematical structure and between learned knowledge and context situations (see Treffers, 1993; Wubbels, Korthagen & Broekman, 1997). Since in English and in other languages the translation of the term 'realistic' will be associated with 'reality' there were attempts to clarify how reality and realistic should be defined in mathematics education settings (Greer, 1997; Säljö, 1991a, 1991b). As van den Heuvel-Panhuizen (2001a) emphasizes the original Dutch term 'zich realizeren' means 'to imagine', therefore realistic mathematics does not always has the real world as context for tasks; objects of the fantasy world (which can be imagined, represented, and therefore modeled) can form an equally appropriate context for mathematization. The current interpretation of the term 'realistic' is a reference to what is *experientially* real (Gravemeijer & Terwel, 2000; Linchevski & Williams, 1999), declaring that not every everyday-life problem will be necessarily experientially real for the students.

Even though there are signs that there was greater emphasis on links to reality fifteen years ago than there is now in the research and development work of RME (see van den Heuvel-Panhuizen, 2000), the strong and relevant connections between real-life contexts and students' mathematical learning is still a major characteristic of RME. Treffers (1993) developed the concepts of horizontal and vertical mathematization. The term mathematization was developed by Freudenthal (see van den Heuvel-Panhuizen, 1996, 2000, 2001a, 2001b, 2003). Mathematization refers to the processes of mathematical activity; since it is not mathematics as a closed system that should be taught in school, but rather the activity of organizing matter from reality. Treffers' horizontal mathematization concept refers to the process of bringing mathematical tools forward in order to organize and solve daily life problems. Vertical mathematization refers to inner mental reorganization of concepts and operations within the mathematical system. Horizontal and vertical mathematization processes are intertwined in students' mathematical activities, and mathematization "contains, in fact, all of the important aspects of the RME educational theory" (van den Heuvel-Panhuizen, 1996, p. 11.).

One crucial point in RME is introducing mathematical models (in a very broad sense of this word). Creating and developing models *for* problem situations is very different from searching for models *of* problem situations (see van den Heuvel-Panhuizen, 2001a). Effective use of several models in different age-groups and in different content areas has been evidenced. Gravemeijer (1994) investigated the empty number line as a powerful mathematical model for several reasons. By means of visualization it enables for using and explaining various strategies, e.g., subtracting 49 can be substituted by subtracting 50 and adding one, or in case of subtracting a relatively large number (e.g., 51 – 49) it may be easier to step forward from the smaller quantity to the larger quantity.

Klein, Beishuizen and Treffers (1998) added that it is not the empty number line alone that contributes to the success of their development program, but the way it was used, i.e. stimulating and discussing different solution patterns in a positive classroom climate. Keijzer and Terwel (2003) studied the understanding of fractions, and also successfully used the number line model (also by means of computer games) to develop understanding. Doorman and Gravemeijer (2009) conducted an experiment among 10th grade students in the field of velocity problems, using discrete graphs as models for reasoning about the relation between displacement in time intervals and total distance traveled. An extension of the RME principles to higher school grades had been previously demonstrated by Gravemeijer and Doorman, (1999) in the field of calculus. In that case determining velocity from time/interval graphs became a model for reasoning about integrating and differentiating arbitrary functions. Van Garderen (2007) argues that diagrams as mathematical models provide the flexibility for children with learning disabilities to generalize what they have learnt in a given situation to another situation.

The realistic mathematics approach proved to be useful also for low attaining students. The principles and suggestions concerning RME for low attaining students have been reviewed by Barnes (2005). Low attaining students and even special education need students profited more from so-called guided instruction, i.e., when much more space is provided for individual contributions, than from a so-called structured or direct instructional approach (Kroesbergen & van Luit, 2002). However, in general, the relationship between mathematical instructional approaches (namely, traditional and realistic approaches) and mathematical proficiency has not been unequivocally evidenced. In general, there are larger differences in pupil performance within a particular mathematics instructional approach than between two different approaches (Koninklijke Nederlandse Akademie van Wetenschappen, 2009).

The Curricular Shaping of Mathematical Literacy

Scientific discourse on the role and importance of curricular aims and objectives has recently been permeated by a range of different curricula as defined according to different levels or phases of the teaching-learning process. When analyzing research-based curriculum development, Clements (2008) narrows the term to available curriculum, i.e. curriculum for which teaching materials exist. There is a usual trinity of curriculum terms used in the (mathematics) education literature: declared, implemented and achieved curriculum. The declared curriculum refers to educational documents set out in different levels of the educational system: national core curriculum, local curricula etc. The implemented curriculum refers to the processes actually carried out in schools, and achieved curriculum refers to students' performance on tests measuring curricular objectives.

In Stein, Remillard and Smith (2007), a diagram shows the relationships between curriculum-related variables including student learning. Although the sequence of the above-mentioned three curricular concepts is straightforward, how these concepts can transform into each other can be explained by several factors. Figure 2.1 also points to the complexity of factors explaining the transition between curricular concepts, listing mutually and necessarily intertwined phenomena as teachers' beliefs, teachers' professional identity, and higher system-level variables as organizational and policy aspects.

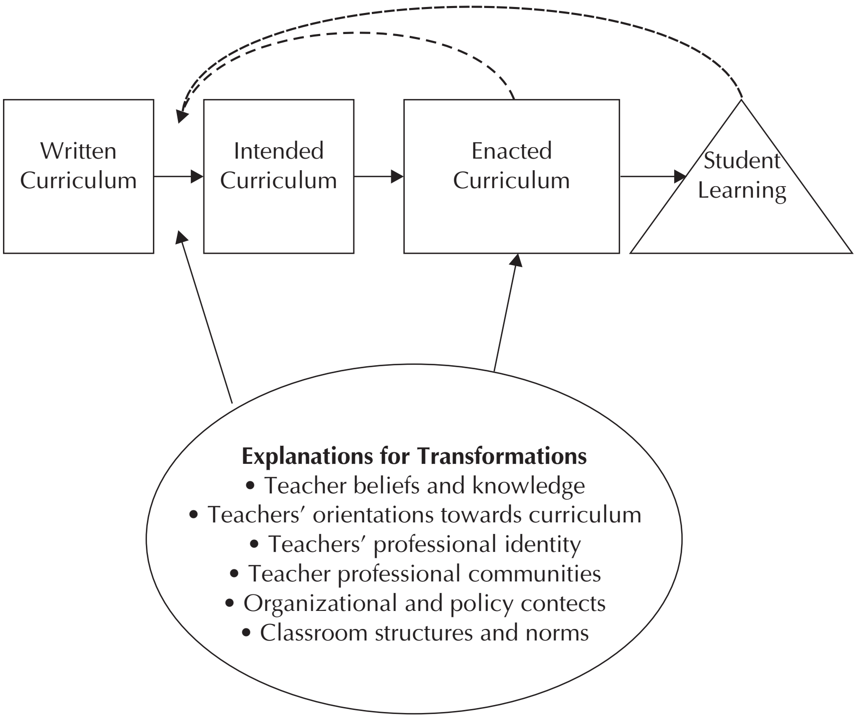


Figure 2.1 Relationships between written, intended and enacted curricula, and student learning (Source: Stein, Remillard & Stein, 2007, p. 322)

Some mathematics task-related factors concerning curricular shaping of knowledge are discussed in Henningsen and Stein's (1997) study. There are at least two steps in between the tasks formed on the basis of the declared curriculum and students' learning outcomes (i.e., the achieved curriculum). Mathematics tasks are set up by the teachers according to their implemented curriculum, and mathematics tasks are in a further step implemented by the students in the classroom. The transition between teacher and student implementations as mentioned in the previous sentence is influenced by several factors including general classroom norms and content-specific sociomathematical norms (Yackel & Cobb, 1996), and teachers' instructional dispositions. The importance of teachers' beliefs and instructional dispositions will be illustrated in the chapter part entitled "tasks measuring mathematical literacy in the classroom".

In this section we focus on examples from national (declared) curricula, since in some way or another, through several direct and indirect factors, national curricula have their impact on both implemented and achieved curricula. The following examples express how in the last decades our curricula declared and emphasized the importance of approaching classroom-based mathematical knowledge and the mathematical knowledge that is transferable to different types of problems and to other school subjects.

Characteristics of Core Curricula in Mathematics

Before introducing the current National Core Curriculum, the so called "National Curriculum '78" had great impact on the Hungarian school system not only because of its descriptive nature (this national curriculum was compulsory for every schools and there were no local curricula) but to the progressive changes it introduced – among others in the field of mathematics. The mathematics part of the national curriculum followed the structure of other parts of the curriculum, i.e. there were aims, objectives and contents formulated for grades 1–4 and grades 5–8, but C. Neményi, Radnainé and Varga (1981) defined overarching intervals for curricular objectives: the divisions of grades 1–3 and grades 4–5 expressed their beliefs that the necessarily continuous developmental processes in students' mathematical thinking should not be separated into two formally distinct stages at the end of the fourth grade (which is a formal dividing line in Hungarian educational system between lower and upper grades of the primary school).

Among the general objectives of the National Curriculum '78 we found motivation in the sense that students are expected to be interested in, and be fond of mathematics both because of external reasons like utility and applicability and because of internal reasons like harmony, truth and beauty in mathematics. (p. 262). According to Aiken (1970), attitudes towards mathematics in adulthood are determined by childhood experiences, and grades 4 to 6 are of crucial importance in forming attitudes. In Hungary, a nationwide analysis revealed that students' attitudes towards mathematics are of mediocre level (Csapó, 2000).

Other curricular objectives present in the National Curriculum '78 pay special attention to student characteristics of a cognitive nature. As for the application of mathematical knowledge in different context, the following objectives were formulated.

In grade 4 and 5 "judgments about (discussion and defending of) unambiguity of tasks, whether a task contain redundant data, incoherent conditions, and whether a given solution process is suitable." (p. 262.) Among the more concrete objective that are connected to a given grade, in grade 5, we found "ability to determine what data are redundant, and what data should be presented in a word problem", an objective that usually (albeit implicitly) implies horizontal mathematization processes. By the end of grade 3, students are required to "be proficient in gathering and organizing data of a word problem. Students must be able to find an appropriate mathematical model (drawing, displaying, operations, open statements), and to solve a word problem by means of that model or by means of trial and error" (p. 283.) The latter objective more explicitly refers to the need of horizontal mathematization in word problem solving.

The National Core Curriculum (Nemzeti alaptanterv; first version: 1995, latest version: 2007) leaves more space for school autonomy, and formulates nationwide curricular objective more loosely and more generally. It is the local curricula that have to elaborate the general nationwide curricular aims and objectives. In line with current trends in international system-level survey requirements, the definition of 'mathematical competence' contain as important element that "the individual is able to apply basic mathematical principles and processes in acquiring knowledge and in solving problems in daily life, at home and at the workplace." (p. 9.) Most of the age-related objectives in the National Core Curriculum are attached to more than one – two year long each – age intervals.

The structure of the NCC objectives follows the two year long interval scheme, i.e. the first milestone in objectives is the end of the second grade, the second milestone is the end of fourth grade etc. The second aspect of the curricular objectives in NCC is the sub-domains of mathematical literacy. One of the sub-domains is labeled as "Application of knowledge". This sub-domain contains curricular objectives explicitly referring to daily life situations and other school subjects. The objective of applying mathematical knowledge in daily life situations is prescribed from the third age cohort (i.e., from grade 5) to grade 12 throughout all grades. The current evaluation framework may and should address the importance of this objective from as early as the first grade of schooling. The relation between knowledge acquired in the classroom and possible applications in real life situations should be strengthen by means of both instructional and evaluation methods.

As Hiebert et al. (1996, p. 14.) warns, "the tension between acquiring knowledge and applying it is not special to mathematics". "The separation of school learning from 'everyday life' has become a problem receiving significant attention by researchers focusing on the sociocultural nature of cognition" (Säljö, 1991a). However, according to Hiebert et al., an emphasis on the application dimension of knowledge may result in less predictable curricula and teachers may worry about the loss of important information, i.e. not covering some parts of the curriculum because of working with time consuming application tasks. The characteristics and problems of math teacher education cannot systematically be reviewed here, albeit some features are highlighted by Szendrei (2007) who reviewed tendencies and efforts in Hungarian mathematics education and mathematics teachers education research from 1970. One of her most important suggestions is that in math teacher training more time should be dedicated to the didactics of mathematics – currently much stronger emphasis is put on the teaching of mathematics itself.

Applications of and Demand on Mathematical Knowledge in other School Subjects

Historically, mathematics fulfilled a leading role in the development of sciences. As Maddy (2008) expresses, till the seventeenth century, great thinkers of those times could not separate mathematics and science. It was the nineteenth century when mathematicians began to develop concepts that had no direct physical meaning. The historical development of mathematics and sciences still has its effects on school curricula and on classroom practice. Interestingly, the Hungarian National Core Curriculum (Nemzeti Alaptanterv, 2007) does not explicitly mention the terms mathematics or mathematical when detailing the learning objectives of the cultural domain "Man and nature". However, within the cultural domain "Our Earth and environment", there are several points in which the role of mathematical abilities (competencies) in geographical and environmental knowledge acquisition is emphasized. There are three main clusters described in which the importance and role of mathematics can be understood: (1) numerical skills for measurements and data handling, (2) spatial intelligence for spatial orientation and (3) logical reasoning, especially in understanding complex spatial and environmental systems.

In sum, there are unexpectedly few explicit relations between mathematics and science objectives in the Hungarian NCC. Of course, there are connections made by teachers between science topics and mathematical prerequisite knowledge, but Pollak's (1969, p. 401) older critical comment that "the student is typically not given the opportunity to participate in making the abstraction from the physical reality to the mathematical model" still applies to the current classroom practice. Some changes are expected to appear in the near future, in part due to the Rocard-report (High Level Group on Science Education, 2007) on inquiry-based learning and the projects just have started like PRIMAS (Promoting Inquiry in Mathematics and Science Education).

The Definition of Mathematical Literacy in the PISA Studies

The PISA (Programme for International Student Assessment) studies aim at defining and measuring students' knowledge and skills in important areas as mathematical, reading and scientific literacy. It was the PISA 2003 study that focused on mathematical literacy (OECD, 2004). This document emphasizes that the "literacy approach" expresses the intention to define and assess mathematical knowledge and skills not in terms of mastery of the school curriculum, but in terms of readiness for full participation in society.

Based on the more general economic definition of "human capital", the PISA studies define mathematical literacy as follows (OECD, 2003, p. 24):

> "Mathematical literacy is an individual's capacity to identify and understand the role that mathematics plays in the world, to make well-founded judgments and to use and engage with mathematics in ways that meet the needs of that individual's life as a constructive, concerned and reflective citizen."

The components of this definition are further elaborated in the above-mentioned document, e.g., the term "world" refers to natural, social and cultural objects, and it is further clarified by referring to Freudenthal's oeuvre. The system of the PISA mathematical tasks is based on the above definition of mathematical literacy. Students have to solve tasks belonging to different content, process and context dimensions. Consequently, the criterion of "use and engagement with mathematics" points to the need of mastering mathematical knowledge applicable in different content domains, on different competency levels and in different contexts. The term "reflectivity" calls forth building awareness and meta-representations fostering knowledge transfer processes across domains (Adey et al., 2007).

The importance of the PISA studies and the further possibility of using their results in evidence-based policy making has been convincingly evidenced by several secondary analyses (e.g., see Baumert et al., 2009).

Tasks Measuring Mathematical Literacy

In this section we analyze how classroom tasks of mathematical literacy are used and what characteristics they have. From an educational evaluation point of view, tasks of formative evaluation will be discussed, i.e. tasks that are embedded in the teaching-learning process in order to develop students' mathematical understanding. We focus on tasks of mathematical literacy where the definition of mathematical literacy is taken from the PISA studies. With regard to the application-related objectives of mathematical knowledge, the context dimension of PISA can be understood as the application of mathematical knowledge in different situations (OECD, 2006).

The PISA literacy approach (OECD, 1999) requires students "be involved in the full mathematical modeling cycle" (Palm, 2009, p. 3), solving tasks that address even out-of-school settings. Although the PISA mathematical literacy has been worked out for measuring 15 year old students' achievement, as we would like to emphasize, even young children's mathematical literacy can be improved and measured in different contexts, in different fields of application.

Characteristics of Classroom Mathematics Word Problems

In this section we restrict our analysis of mathematical tasks that are relevant from the aspect of application of mathematical knowledge. Since the application of mathematical knowledge usually requires the use of textual elaboration (at least in the phase of posing the problem), word problems will be in the focus of our analysis.

"Word problems can be defined as verbal descriptions of problem situations wherein on or more questions are raised the answer to which can be obtained by the application of mathematical operations to numerical data available in the problem statement." (Verschaffel, Greer & De Corte, 2000, p. ix.)

Historically, word problems fulfilled two interfering roles during the last several centuries. From as early as ancient river valley civilization times, mathematical word problems provided the means for mastering arithmetical skills and at the same time providing tools for solving daily life problems that were of crucial importance in a certain historical context. Work of ancient Egyptian workers or computations necessary to be a successful Venetian merchant required both high-level arithmetical skills and strong connections between problems arisen from daily life and between mathematical prototype examples (see Verschaffel, Greer & De Corte, 2000). This duality of the functions of word problems has lived on till today, and the interference and the state of being intertwined result in questions about the effective use of word problems in classrooms.

The importance of word problems in improving the applications of mathematics has been justified by Pollak (1969, p. 393) in the following way: "How does the student become involved in applications of mathematics? Throughout most of his education, mainly through … 'word' problems".

Types of classroom mathematical word problems may be grouped and analyzed according to textual, semantic and mathematical features they have. Educated people can easily distinguish among different types of word problems. As Säljö (1991b) pointed out, even the twentieth century reader can easily recognize the genre of a mathematical word problem text, and may be capable to handle texts like the following one from 1478:

> *If 17 men build 4 houses in 9 days, how many days will it take 20 men to build 5 houses?*

As long as the solver knows that there exist a direct proportional relation between the number of men at work and the number of houses being built, "our familiarity with this genre leads us to recognize that the extra-linguistic activity that is being referred to – building houses – is, if not accidental, at least not central to the task as an exercise in elementary arithmetic." (Säljö, 1991b) The content of this task can be varied without restraint, and it is not necessary to know any house-building technologies or team working method to solve the task. What is more, it would be disadvantageous to start a deep semantic analysis of the reality of task variables. “The pseudo-real contexts … encourage students to see school mathematics as a strange and mysterious language” (Boaler, 1994, p. 554.). The micro-worlds of word problems (this term is borrowed from Lave, 1992) belong to the same genre of texts, a genre that was caricatured two centuries ago by Flaubert writing his letter about the ill-famed ‘How old is the captain?’ problem.

Boaler (1994) criticized the so-called pseudo-real type of mathematics word problems from a feminist point of view. Although many tasks are equally strange for both boys and girls, in Boaler’s research girls suffered more from pseudo-real context tasks in traditional learning environments than boys. In her own intervention studies, this traditional approach for ignoring the role of content is seriously challenged and uncovered. The main problem concerning the context of school mathematics word problems is suspending reality and ignoring common sense due to entering the genre of word problem texts. According to Boaler (1994), this difficulty can be overcome by changing instructional methods towards a process-based learning environment. Process-based learning environments, where all students work on open-ended problems and are encouraged to investigate and to discover mathematics, proved to lessen sex differences in mathematical achievement (see also Boaler, 2009).

Classroom mathematics word problems may have another facet that hinders students’ development. In the field of learning fractions, Mack (1990) has revealed that the sequence of tasks does not correspond to the sequence how students’ prior knowledge would help understanding fractions. Concretely, six grade student have ample prior experience about fractions, and they often use partitioning (i.e., dividing quantities into pieces), and thus they can relatively easily understand improper fractions (i.e., when the numerator is greater than the denominator). However, tasks containing improper fractions are usually left to the end of the fraction chapters in the textbooks.

A similar problem has been found with multiplication by Lampert (1986). She emphasizes that in students’ mind multiplication is more complex than repeated addition. If we limit though instruction one’s mental model about multiplication to additive compositions, the student may fail later in understanding multiplications to continuous quantities. Lampert’s and Mack’s research results nicely support more general recent principles of mathematics education like the RME mathematization concept. Schoenfeld's (1988) heretic standpoint about the disaster of well taught lessons tells the same story: carefully performed sequence of steps in constructing mathematics gives the message to students that it is the (mathematical) accuracy that counts when doing mathematics. How students' experiences can provide unexpected results in mathematical word problems were documented in research on child street vendors (Carraher, Carraher & Schliemann, 1985; Saxe, 1988). Although from a mathematical aspect larger natural numbers are more difficult to add and subtract, children having experiences with the inflated Brazilian currency were better in adding numbers that could be matched with real prices even if these numbers were relatively large.

Classroom word problems were categorized in several investigations according to features that are both mathematical and of cognitive representation nature. As far as additive structures are concerned, the following types of simple word problems were identified: combine, compare, change and equalize problems (see Radatz, 1983; Riley & Greeno, 1998; Jitendra, Griffin, Deatline-Buchman & Sczesniak, 2007; Morales, Shute & Pellegrino, 1985).

Independently of the task content, students strive for categorizing word problems, and driven by their beliefs about the solvability of word problems, form different strategies to cope with different types of problems. This tendency to categorize problems is not per se a problem, since recognizing the common structure of superficially varying tasks is an important characteristic of true expertise in a given domain (see e.g., Sternberg & Frensch, 1992). However, when finding the operation to be computed and the data to be matched with that operation are generally sufficient for solving a task, it may create blind alleys for students in their mathematical development. Verschaffel, Greer and De Corte (2000) analyze this so-called superficial schema of word problem solving, comparing it to the schema of genuine mathematical modeling. The crucial point is whether the student builds a situation model by means of deep understanding of the problem situation, or (s)he skips building such a situation model and jumps immediately to a mathematical model deemed to be appropriate – based on superficial task characteristics. Illustrating and documenting those blind alleys in word problem solving the reader should consult Verschaffel, Greer and De Corte (2000). A Hungarian study brought further evidence about the presence and strength of superficial word problem solving strategies (Csíkos, 2003).

One important aspect of using word problems in classrooms is teachers' beliefs and attitudes towards realistic word problems. "The teachers seem to believe that the activation of realistic context-based considerations should *not* be stimulated but rather discouraged in elementary school mathematics" (Gravemeijer, 1997, p. 391. – italics in original text). Verschaffel, De Corte and Borghart (1997) empirically documented pre-service teachers' disposition towards giving non-realistic reactions to simple arithmetic word problems themselves as well as their tendency to give higher marks to non-realistic than to realistic interpretations and solutions of word problems by students.

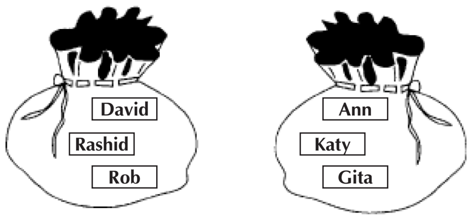
Sociomathematical Norms, Contextual and Content Effects

The term "sociomathematical norms" was introduced by Yackel and Cobb (1996). These norms, which are (in contrast to the broader social norms) by definition restricted to the curricular domain of mathematics, are derived from individual and group mathematical activities (classroom practices). Classroom teachers as representatives of the mathematical community (Yackel and Cobb's experiment was carried out in second grade classrooms) have a crucial role in establishing norms about mathematics and its teaching and learning like what an appropriate mathematical problem is, what an appropriate response to a mathematical task is, how the acceptable forms of explanation and argumentation look like, etc. These norms can vary from classroom to classroom, but "sociomathematical norms are established in all classrooms regardless of instructional tradition" (p. 462).

One important aspect of sociomathematical norms is whether acceptable mathematical explanations in a classroom are mathematical or status-based. Many children tend to infer that their answer is incorrect as soon as the teacher questions it. This norm can easily lead to rigid and false beliefs about the nature of mathematical problem solving and argumentation. Although the analysis of children's mathematical beliefs is beyond the scope of this chapter, it is students' mathematical beliefs that take their share in explaining difficulties in the application of their mathematical knowledge in different contexts and settings (e.g. in mathematics in streets versus in schools, see Carraher et al., 1985). One strong belief revealed in several studies is that a mathematical task always has (only) one right solution, and there is (only) one right way to find that solution (see for example Reusser & Stebler, 1997; Verschaffel, Greer & De Corte, 2000; Wyndhamn & Säljö, 1997).

How sociomathematical norms in general and norms about the role of reality in word problem solving in particular develop can be understood in the light of some theories belonging to sociology and linguistics. Cooper (1994) has successfully used Bernstein's educational knowledge codes, distinguishing between common sense knowledge and school knowledge (also called everyday and esoteric knowledge, respectively). According to Bernstein's argument, children are very early in their school career discouraged from connecting common sense knowledge and school knowledge. Even today it can be revealed that school success depends to some extent on students' willingness and capacity to disclose common sense knowledge as a source of information in mathematics problem solving. Cooper and Dunne (1998) applied both Bernstein's and Bourdieu's insights about the possible social class differences in school (and mathematics) achievement. These differences can be attributed to a relative lack of access to the cultural resources demanded in school situations. Bourdieu's powerful phenomenon of "feel for the game" could be applied in explaining social class differences in some standardized mathematics items. One striking example is the so-called Tennis item depicted in Figure 2.2.

David and Gita's group organize a mixed double tennis competition. They need to pair a boy with a girl. The put the three boys' names into one bag and all the three girls' names into another bag.



Find all the possible ways that boys and girls can be paired.
***Write the pairs below**. One pair is already shown.*

Rob and Katy
...

Figure 2.2 The Tennis item. Source: Cooper and Dunne, 1998, p. 132.

Detailed analyses of students' achievements and interview transcripts have shown how the "feel for the game" phenomenon explains social class differences. For esoteric mathematical reasoning, it is clear that children's names and supposed nationality is not a relevant consideration to be taken account of. About one quarter of students aged 10–11 years produced only three pairs instead of the mathematically correct nine ones. However, these children produced three "realistic" pairs in a sense that the three pairs were distinct; each name was used only once. According to Cooper and Dunne, this type of tasks used in evaluation settings raises problems of equity, i.e. equal opportunities in education. How in general mathematics word problems generate inequities (in terms of gender, social class, etc.) is analyzed and criticized also by Boaler (2009).

According to other empirical results, in grade 3, word problems of the story problem type (i.e., where figures and relations are embedded in a narrative story) are challenging for students (Jitendra, Griffin, Deatline-Buchman & Sczesniak, 2007). Nevertheless, in grade 3, word problem solving is a useful indicator of general mathematical proficiency (Jitendra, Sczesniak & Deatline-Buchman, 2005)

The role of culture in mathematics achievement incorporates the role of language competence. To understand mathematical word problems one has to be capable semantically analyze the linguistic components of a task, and furthermore, to identify important and redundant parts. Elbers and de Haan (2005) studied multicultural classrooms in which language components of mathematical word problems are of more peculiar importance. They found that language problems in understanding texts were not solved by means of referring to the everyday meaning of words, but conversations (and students' help-seeking behavior) focused on the special meaning of terms they have in the context of a mathematical lesson. The priority of understanding word problem text genre and context over pure semantic understanding of text cues have been further supported by Morales, Shute and Pellegrino (1985) whose study revealed no language effect on either solution accuracy or on the ability of categorizing math word problems – their subjects were Mexican-American. Nevertheless, well-documented results prove that the linguistic features of a word problem influence to certain extent the solution process (e.g. the term 'of these' may influence whether an appropriate mental representation is built, see Kintsch, 1985).

Two effective strategies to promote connections between students' mental representations and learning objectives to meet can be: rewording the word problem, or personalizing it. In an investigation by Davis-Dorsey, Ross and Morrison (1991) it has been revealed that fifth grade students profited from the personalization of the task (i.e., incorporating personal information about the learner) and second grade students profited from both personalizing and rewording the content (i.e., making the text more explicit, helping to translate its content into mathematical terms). In this experiment, word problems that could be considered as mathematically identical, did differ in their contextual and content features.

Another – even more radical – possible change in improving classroom environment is the use of reciprocal teaching in mathematics. Magdalene Lampert (1990) adapted the instructional method called reciprocal teaching from reading education (see also van Garderen, 2004). The heart of this method is deliberately altering the roles and responsibilities of the teachers and students in the classroom. She notes that this change requires changes also in tasks that define mathematical lessons. As for defining different contexts in which the application of mathematical knowledge is claimed and expected, we follow Light and Butterworth (1992) who gave a rather broad definition: the context of a task consists of several layers of information related to the task: physical, social and cultural settings. Tasks with the same mathematical structure and with the same content can be solved differently according to changes in the context. However, as Verschaffel, Greer and De Corte (2000) illustrate, the effects of context changes, in case of a special class of word problems context changes, may result in only slightly different levels of student achievement. These context changes involved warning messages at the top the paper and pencil tests or embedding the task in a test that contain puzzle type tasks. These slight changes may suggest that context changes more radical than staying within the paper and pencil methodology may have stronger influence on students' solution patterns.

The content of a task can be defined as taking the definition of context as a starting point. We also borrow the expression 'noun term' from Kintsch and Greeno's (1985) seminal article. There is an assumption widely accepted (or at least used) in the mathematics education community: word problems should fulfill the role of providing a parade-ground for mastering arithmetic skills. According to this tradition, changing the content of a task should not necessarily influence students' achievement; what is more, students are expected to develop transfer skills enabling them to solve tasks with the same deep mathematical structure equally well, independently of the current content elements of he tasks. It should make no odds whether the noun terms of a task originate in the micro-worlds of football or fashion or whether some superficial changes are made in the formulation or the placement of the givens and/or the question.

A Taxonomy of Tasks of Mathematical Literacy

In this section a categorization of mathematical tasks will be proposed. There are many aspects that can be starting points for different categorizations. In international system-level surveys (see e.g., OECD, 1999) there is usually a multidimensional model in which tasks are classified according to mathematical content, thinking processes required, and task format. In the PISA studies (see OECD, 2003) the context of the task appeared as a new dimension. The existence of the context dimension and the four values of this scale can be considered as an expression of an educational policy intention of paying ample attention at the applied side of mathematics and of covering a wide range of topics in assessing mathematics literacy.

When applying two or three dimensions (e.g. mathematical content, context, and competency cluster in PISA 2003) and the concrete values of each dimension, a rectangle or cuboid can be used as a model of which there are several cells representing different types of tasks. Now we provide a category-system for an 'application' dimension of mathematical knowledge. This categorization has its precedents in part in the PISA study contextual dimension, but mainly relies on the horizontal mathematization idea of the RME movement.

Challenges and Difficulties in Developing a Category System for Application Tasks

The logic and basis for this categorization is in line with Erikson's (2008) idea of developmental stages in arithmetical thinking. Different developmental stages can be associated with corresponding behavioral patterns and corresponding mental structures. Starting from a possible hierarchy of mental structures, it is possible to match them with corresponding behavioral patterns observable in appropriate evaluation contexts. In this sense, tasks unambiguously belonging to different categories of tasks requiring different behavioral patterns will make it possible to reveal the test takers' corresponding mental structures. However, with respect to the application dimension of mathematical knowledge, there are problems with matching mental processes and observable behavior. A striking example came from Cooper (1994). The so-called Lift problem (Figure 2.3) have become an often cited example illustrating how different possible solutions to an open-ended question can be analyzed in terms of understanding the task as a realistic or routine task.

This is the sign in a lift at an office block:

This lift can carry up to
14 people

In the morning rush, 269 people want to go up in this lift.
How many times must it go up?

Fig. 2.3 The Lift problem

In Cooper's (1994) analysis it is clear that the expected right answer (i.e. 269 ÷ 14 rounded up to the nearest whole number can be the result of very different understandings and solution strategies. One possible way is to understand that this task signifies a real problem that has to be solved, but taking account of the test condition, students should not create new variables and should not question some axioms implicitly involved in the task. The other way is to understand that this task signifies a routine school mathematics problem but there is a trap in it. In this second way, one should not divide 269 by 14, because of falling to trap. However, as Cooper suggests, the first type right solution requires some assumptions that are almost never true, e.g. the lift is always full except for the last trip. If someone assumes that a lift that is designed for 14 people works on average carrying about 10 persons, will give a wrong answer if only she realizes that in a test one is not expected to create new variables, but to find out the intentions and use the rules such tasks usually require and activate.

There are some classifications of realistic (and non-realistic) mathematics word problems proposed in the literature. One relevant aspect is whether the task classification has a mental representational and instructional focus or whether it has a system-level assessment purpose. The first aspect is representative of a taxonomy proposed by Galbraith and Stillman (2001). According to Verschaffel (2006), this categorization focuses on student thinking processes expected to elicit and on the relationship between word problems and the real world. In this taxonomy, there are four word problem categories:

(1) injudicious problems, wherein realistic constraints are seriously violated;
(2) context-separable problems, wherein the context plays no real role in the solution and can be stripped away to expose a purely mathematical question;
(3) standard application problems, where the necessary mathematics is context-related and the situation is realistic, but where the procedure is (still) rather standard;
(4) genuine modeling problems, in which no mathematics as such appears in the problem statement, and where the demarcation and formulation of the problem, in mathematical terms, must be (at least partly) supplied by the modeler.

This taxonomy focuses on students' thinking (modeling) processes, i.e. how links between their mental representations and the real-world objects are realized.

Another categorization that can also be considered as an important antecedent of the categories proposed in the forthcoming parts of this chapter, was described by Palm (2008, 2009). Palm focuses on task characteristics of word problems that emulate out-of-school situations. He attempts to describe what characteristics a so-called authentic task should have. The key idea is a reference to the elements of 'simulation', i.e. the concordance between word problems and out-of-school, real-world task situations: comprehensiveness, fidelity and representativeness. These terms are borrowed from a seminal work written by Fitzpatrick and Morrison (1971), whose work was made of a system-level evaluation purpose.

Palm's approach for categorizing authentic tasks yielded support from an analysis of Finnish and Swedish national assessment tasks. Although this task battery was made for upper secondary school students, there are some lessons worth considering for lower grades as well. It has been revealed that 50% of the word problems used in national assessment both described an event that might occur out of school context and included a question that might be ‘realistically’ posed in that event. These two superficial task characteristics may strongly indicate that the word problem is authentic, and authenticity – as described in other taxonomies – is associated with students’ genuine mathematical modeling processes.

Our attempt to set up a taxonomy for word problems from the aspect of applied mathematical knowledge will necessary take account of both characteristics of word problems and the mental processes that are elicited in the word problem solving process. There will be four task categories proposed in a way that it may be considered a two by two system. There are two categories for word problems not requiring genuine mathematical modeling of the problem situation, and there are two categories called realistic and authentic that refer to genuine mathematical modeling in the sense of the following description: In accordance with Galbraith and Stillman (2001), genuine modeling problems are problems wherein there is at least one modeling complexity involved that makes that the solver cannot straightforwardly formulate, understand, mathematically represent, solve, interpret, answer the problem in the same way as he can do for a prototype or pseudo-real problem.

“Bare Tasks” Containing Purely Mathematical Symbols

The term “bare tasks” is borrowed from Berends and van Lieshout’s (2009) taxonomy for word problems in relation with whether they contain drawings as essential or irrelevant part of the task. Bare tasks contain purely mathematical symbols and at most a formal instruction about what to do or how to solve the task (e.g., “$10 + 26 = ?$”). This category stands here as a sufficient and necessary starting point to define what types of tasks have little to do with the application of mathematics. Tasks containing purely mathematical symbols – or text at most ‘solve the equation’ type instructions – do not usually have relations with students’ applied problem solving or mathematical modeling. Please note, however, that even bare tasks are appropriate means for facilitating mathematical modeling in a way that is called a reverse way of word problem solving, i. e. when students are taught how to pose word problems given the mathematical structure of the task in purely symbols.

This type of tasks is usually part of everyday classroom practice, and the capability to solve such tasks is part of the curricular objectives as well. A possible sharp distinction between these 'bare tasks' and tasks of the other three categories can be found in understanding and learning fractions (Mack, 1990).

We do not want to give the impression that bare tasks are per se easier than tasks embedded in a context. To the contrary, in some cases, children will perform better on word problems than on mathematically isomorphic bare tasks. This has been stressed and documented by several authors (Carpenter, Moser, & Bebout, 1988; De Corte & Verschaffel, 1981).

Prototype and Pseudo-Real Word Problems

As we have discussed in a previous section, classroom instruction frequently uses and relies on so-called prototype examples. These tasks are word problems dressed on a skeleton that can be considered as a representative of a mathematical operation or other mathematizing process. Prototype examples are often called in Hungary 'green stove' or 'precept' examples from which one can induce and explore analogies. We define prototype examples as mathematical word problems that are used in order to learn to recognize and practice a particular mathematical operation (e.g. multiplication) or a particular mathematical formula or solution schema (e.g. the "rule of three"), In such problems, the content is carefully selected or constructed because of its familiar and prototypical nature, but that content has no special meaning or role from a realistic point of view.

Certainly, learning form worked-out prototype examples can be a powerful tool in improving students' mathematical abilities, but there is a potential danger in generating so-called rational errors (Ben-Zeev, 1995) in a way that instead of transferring the deep structure and the solution processes adequate for the prototype example students may rely on surface similarities. (E.g., poor learners may categorize word problems according to their content or contextual features like 'age difference tasks', 'flag coloring tasks' and so on even though mathematically speaking they have little or nothing in common.)

The understanding and solving of many word problems depends on "tacitly agreed rules of interpretation and on multiple assumptions of prototypicality" (Greer, 1997, p. 297.) According to Hong (1995), good problem solver sixth grade students are able to categorize word problems in the early phase of problem solving, i.e. already during the initial reading of the problem. Jonassen (2003) provided an extensive review of literature about students' (mis)categorizing word problems. The essence of these studies, as it can be plausibly hypothesized, is that successful problem solvers categorize word problems according to their (mathematical) structural characteristics, while poor achievers tend to rely on surface (or situational) features (see Jonassen, 2003; Verschaffel, De Corte & Lasure, 1994). It is not mainly the content of the task that elicits such superficial strategies, but the feedback received from the teacher (and from other participants of the school system) about the sufficiency of using such strategies. Many teachers even explicitly teach four- or five phase strategies by which most of the word problems can be successfully solved (e.g., gathering the relevant data, naming the necessary operation, executing the operation, underlining the solution) Teaching such strategies is saluted only if the meaningfulness (or mindfulness) and the flexibility (or adaptivity) of these strategies can be maintained.

Realistic Word Problems

The assessment of student achievement on realistic word problems must, however, be done more flexibly and more dynamically than in traditional former ways (Streefland & van den Heuvel-Panhuizen, 1999).

The term 'realistic' is used according to the Dutch RME definition. In a realistic problem, students are expected (and many times required) to use their mental representations and models in order to understand and solve the problem. Please note that the term realistic refers to mental imageries that are the various means for appropriate problem representations. However, activating and using mental imageries do not necessarily imply that a task is realistic. In Cobb's (1995) understanding, adding two two-digit numbers will not require students to use situation-specific imageries, albeit they probably use imageries during the addition process. Making distinction between realistic and pseudo-realistic word problems the term' situation specific imagery can be of our help.

How to distinguish realistic word problems from the prototype- or pseudo-realistic ones? We agree with Hiebert et al. (1996) that no task in itself can be routine or problematic. A task becomes problematic to the extent and by means of treating them problematic. Likewise, a word problem becomes realistic to the extent it enables students to use their mental images based on real-world experiences. Inoue (2008) suggests helping students validate problem solving in terms of their everyday experiences. It can be done by incorporating fewer contextual constraints in order to let students create a richer opportunity for imaginary construction of the problem. This is in line with Reusser's (1988) observation, who found the various textual and contextual cues too helpful in anticipating the problem solving process. For example, students too often think they are on the right way if the solution process works out evenly (e.g., a division can be executed without a remainder).

In many cases, realistic word problems usually have relatively longer texts than prototype or pseudo-realistic problems do. This is justified by Larsen and Zandieh (2008) in the case of algebra items, where they found it necessary to have a wordy explanation of the situation – when the item is situated in a realistic context. Consequently, the length of the problem text in itself is not a criterion.

A general criterion of a word problem being realistic will involve the following criterion: In a given age-group, for the *majority of students, solution requires mental processes involving horizontal mathematization and genuine modeling elements that go beyond the mere application of a previously taught and well-learnt operation, solution scheme or method. Realistic word problems enable student to build different mental models of a problem situation.* These models may range from mental number lines to a sketched drawing of a rectangular.

Let us illustrate the functioning of this criterion with a task posed by Gravemeijer (1997):

> *Marco asks his mother if his friend Pim may stay for dinner. His mother agrees, but this means that there is one cheeseburger short. There are five cheeseburgers, and including Pim there are six people now.*
>
> *How would you divide five cheeseburgers between six people?*

As Gravemeijer notes, in a real life situation, there can be different practical solutions given: e.g., Marco shares his cheeseburger with his friend, father and mother share their cheeseburger to help out or someone goes out to buy an extra one. Of course, in the mathematical classroom, where all theories of tasks contexts born in the previous decades tell their own story (“feel for the game”, sociomathematical norms, mathematical beliefs, dual educational codes), hardly anyone will propose a solution similar to the above mentioned three renegade answer except for those who do not feel themselves competent enough in division-like tasks. We may hypothesize that more first and second grade children will give renegade, contextual answers taken account of the situation variables than older children would. As for an upper estimation, hopefully the majority of seventh and eighth grade students is able to compute 5/6 as a result of a division called forth by the text of the problem, and without mobilizing situation-dependent imageries. Consequently, this ‘Cheeseburger item’ might serve as a realistic task in grades 3 to 6, requiring students to activate situation-dependent imageries, and find an appropriate mathematical model for the solution. Furthermore, for older children, the task may appear as a prototypical word problem, since they are able to divide 5 by 6, whatever concrete objects are mentioned in the problem statement.

There are useful considerations proposed in the literature about how a word problem may become realistic. According to Boaler (1994), students often do not see the connections between mathematical situations presented in different contexts, and this is because of the (pseudo-real) contexts used in mathematical classroom. She suggests careful selection and construction of word problems in order to develop transferable knowledge from the classroom the ‘real world’. Mere replication of real life situations in word problems is not appropriate. To clarify the difference between word problems that facilitate students’ knowledge transfer from their real world experiences, the following example may be helpful.

De Lange (1993, p. 151.) cited an example from the Illinois State test:

Kathy has bought 40 c[1] *worth of nuts. June has bought 8 ounces*[2] *of nuts. Which girl bought the most nuts?*

a June
b They both bought the same amount
c Kathy bought twice as much
d Kathy bought one ounce more
e You can’t know

1 c stands for cents, i.e., 40 c equals .4 USD.
2 8 ounces is a half pound, i.e. about 22.7 dkg

According to de Lange, the attempt is „admirable", since solving this problem requires the student to make an appropriate mental model for the situation, and any attempt to use a general strategy like „search for the data, choose the right operation, and execute the computation" would fail. The expected right solution here is "you can't know", since the numerical data will not imply any straightforward computational answer. However, de Lange suggests to further improve the task in a way that all options might be true, and it is the students who have to create different task conditions in which the options become true. Furthermore, it follows that the task format in itself can make a problems situation realistic: often it is the open-endedness of a task that makes a given word problem realistic.

In Treffers' example (1993) the use of newspaper excerpts revealed how children can try to solve without bias a mathematical word problem. Fourth grade children receiving the text saying that "On average I work 220 hours per week" was questioned whether it was possible to work 220 hours per week. Children not immediately mathematized the problem, and give answers of various types. One important aspect of realistic mathematics tasks is to encourage diversity by means of open-endedness.

Contrary to previous assumptions, as Inoue (2008) warns, the benefit of use of familiar situations is limited. What is more, the familiarity of the context seems to be correlated with both the content area within mathematics and with the required level of thinking processes (Sáenz, 2009). For example, open-endedness in question format is more frequently related to higher level thinking skills. – Hence the three dimensions of the mathematical objectives (disciplinary content, applied mathematical knowledge, mathematical thinking abilities) are intertwined, enabling us to consider the application dimension as albeit relatively distinct, but embedded in different category values of the other evaluation dimensions.

Authentic Word Problems

A fourth type of word problems is labeled as authentic. Although it should be clear that the terms realistic and authentic are closely related, we feel the need to use the term authentic word problems to give a specific qualification to a particular subset of realistic word problems. The term 'authentic' has been used in various contexts in the mathematics problems solving literature. Accepting Palm's definition, authenticity has several degrees, and it expresses a relation between school tasks and real life situations. When "a school task …well emulates a real life task situation" (Palm, 2008, p. 40) that task may be called an authentic one. On the other way, Kramarski, Mevarech and Arami (2002) approached authenticity from a problem solving perspective. They call a mathematical task authentic if the solution method is not known in advance or there are no ready-made algorithms. A third proposal for a definition comes from Garcia, Sanchez and Escudero (2007) who speak about authentic activities, i.e. the process of relating a task and a real situation.

In itself no task can be considered either authentic or non-authentic (similarly to the lack of distinction in case of the realistic versus non-realistic dichotomy), so when aiming at providing useful categories for an evaluation framework, these three definitions are not equally applicable. As for the first definition, emulating a real life task situation may refer to two things when making decisions about the level of authenticity. First, the degree of emulation may depend on a textual elaboration or creating an appropriate task context (e.g. playing the situation). Secondly, there can be remarkable differences among students in that to what extent a situation can be of familiar (therefore real life) nature. The second definition has even more obviously addressed inter-individual differences (i.e. a solution method is not known for whom?). The third approach is closer to the RME interpretation of horizontal mathematization. In sum, from educational evaluation purposes, we suggest using Palm's definition with emphasis on the need for extensive verbal elaboration in order to "emulate" real life situations.

From an educational evaluation aspect, characteristics of and requirement for authentic tasks can be summarized along two lines. First, authenticity should usually require an alienation from the traditional individual paper and pencil methodology towards more authentic settings such as group working on tasks consisting of various sources of information. Second, authentic tasks in traditional paper and pencil format will be lengthier in text, since descriptions of intransparent problem spaces will result in longer sentences providing cues for missing information and providing also redundant details emulating real life situations in that way. Furthermore, many authentic task will contain photos, tables, graphs, cartoons etc. What is more, authenticity refers to a kind of task-solving behavior and student activity.

It is worth bearing in mind that reaching authenticity as reflection or emulation of real world events and situations is rather a utopia, since the context of schooling and the context of the real world are fundamentally different (Depaepe, De Corte & Verschaffel, 2009). The so-called realistic and authentic tasks do not always measure mathematical knowledge and its relations to real life situations, but they measure the 'feel for the game' as analyzed in the "Sociomathematical norms…" section. Although the 'feel for the game' is a valuable aspect of one's achievement, the possibility of totally different mental representations resulting in the same (right) answer to a task intended to measure the application of mathematical knowledge in an everyday context, urged Cooper (1994) to warn politicians and researchers in a way that

> *Mathematics Education "the English experience [in evaluating mathematical knowledge in everyday context] so far suggest that both much longer times scales to allow for the lessons of research and experience play a greater role, and less political interference in the development of tests, will be needed" (p. 163.)*

As Hiebert et al. (1996, p. 10) suggested, "problematizing depends more on the student and the culture of the classroom than on the task." A problem that can be a routine task in one classroom can be problematic and require 'reflective inquiry' while „given a different culture, even large-scale real-life situations can be drained of their problematic possibilities. *Tasks are inherently neither problematic nor routine*. (p. 10. – italicized by us).

In sum, authentic tasks usually have the following characteristics:

(1) detailed (often lengthy) description of a problem situation emulating real world events

(2) the solution requires genuine mathematical modeling of the situation

(3) the solution process often requires so-called 'authentic activity', e.g. gathering further data by means of various methods (measuring, estimating, discussing prior knowledge about a topic)

(4) in many cases students are encouraged to pose problems and ask questions based on both the given word problem and on their real-world experiences.

Summary

Even though bare arithmetic tasks and prototypical word problems still deserve a place in elementary school mathematics teaching and assessment, they need to be complemented more than was the case hitherto with other, more realistic and more authentic types of tasks, which have recently shown to be more promising vehicles for realizing the “application function” of word problems, i.e. to offer practice for the quantitative situations of everyday life in which mathematics learners will need what they have learned in their mathematics lessons.

By their very nature, those realistic and authentic problems have a greater potential of providing learning experiences wherein learners are stimulated to jointly use their mathematical knowledge and their knowledge from other curricular domains such as (social) sciences and from the real world, to build meaningful situational and mathematical models and come to senseful solutions. At the same time, these more authentic and realistic problems yield – because of their essentially non-routine, challenging and open nature, ample opportunities for the development of problem solving strategies (heuristics) and metacognitive skills that may – if accompanied with appropriate instructional interventions aimed at decontextualisation and generalisation – transfer to other curricular and out-of-school domains. And they involve many possibilities to contribute at the deconstruction of several inappropriate beliefs about and attitudes towards mathematics and its relation to the real world.

An important but difficult issue for assessment is how to make it clear to the learners what is expected – in terms of the required level of realism and precision – from them in a concrete assessment setting. In principle, the question about the mathematical model’s degree of abstraction and precision should be regarded as a part of what we want students to learn to make deliberate judgments about, as one crucial aspect of a disposition towards realistic mathematical modelling and applied problem solving.

Within the context of a regular mathematics class, wherein discussion and collaboration is allowed and even stimulated, the degree of precision, the reasonableness of plausible assumptions, and so on, may be negotiated (Verschaffel, 2002). But such unclarities and difficulties with respect to the level of realism and precision are more serious, we believe, when problems are presented in a context that precludes discussion, especially an individual written test, as has been shown above when discussing the work of Cooper, 1994; Cooper & Dunne, 1998). So, if we want to include more realistic and authentic problems in our assessments, as pleaded above, we will also need to pay attention at how we will make it clear to the learner – explicitly or implicitly – what "the rules of the game" are for a given assessment problem.

References

Adey, P., Csapó, B., Demetriou, A., Hautamäki, J., & Shayer, M. (2007). Can we intelligent about intelligence? Why education needs the concept of plastic general ability. *Educational Research Review, 2*. 75–97.

Aiken, L. R. (1970). Attitudes towards mathematics. *Review of Educational Research*, *40.* 551–596.

Barnes, H. (2005). The theory of Realistic Mathematics Education as a theoretical framework for teaching low attainers in mathematics. *Pythagoras*, *61*. 42–57.

Baumert, J., Lüdtke, O., Trautwein, U., & Brunner, M. (2009). Large-scale student assessment studies measure the results of processes of knowledge acquisition: Evidence in support of the distinction between intelligence and student achievement. *Educational Research Review, 4*. 165–176.

Ben-Zeev, T. (1995). The nature and origin of rational errors in arithmetic thinking: Induction from examples and prior knowledge. *Cognitive Science*, *19.* 341–376.

Berends, I. E., & van Lieshout, E. C. D. M. (2009). The effect of illustrations in arithmetic problem-solving: Effects of increased cognitive load. *Learning and Instruction*, *19.* 345–353.

Boaler, J. (1994). When do girls prefer football to fashion? A analysis of female underachievement in relation to 'realistic' mathematics context. *British Educational Research Journal*, *20.* 551–564.

Boaler, J. (2009.). Can mathematics problems help with the inequities of the world? In L. Verschaffel, B. Greer, W. Van Dooren, & S. Mukhopadhyay (Eds.), *Words and worlds: Modeling verbal descriptions of situations* (pp. 131–139). Rotterdam: Sense Publications.

C. Neményi, E., Radnainé Szendrei J. & Varga T. (1981): Matematika 5–8 [Mathematics 5-8]. In Szebenyi. P. (Ed.), Az általános iskolai nevelés és oktatás terve [Curriculum for primary school]. 2nd edition. Budapest: Országos Pedagógiai Intézet.

Carpenter, T.P., Moser, J.M., & Bebout, H.C. (1988). Representation of addition and subtraction word problems. *Journal for Research in Mathematics Education*, *19.* 345–357.

Carraher, T.N., Carraher, D.W., & Schliemann, A.D (1985). Mathematics in streets and schools. *British Journal of Developmental Psychology*, *3.* 21–29.

Clements, D. H. (2008). Linking research and curriculum development. In L. D. English (Ed.), *Handbook of International Research in Mathematics Education* (pp. 589–625). 2nd edition. New York and London: Routledge.

Cobb, J. (1995). Cultural tools and mathematical learning: A case study. *Journal for Research in Mathematics Learning*, *26.* 362–385.

Cooper, B. (1994). Authentic testing in mathematics? The boundary between everyday and mathematical knowledge in National Curriculum testing in English Schools. *Assessment in Education: Principles, Policy & Practice*, *1.* 143–166.

Cooper, B., & Dunne, M. (1998). *Sociological Review*, *46*(1). 115–148.

Csapó, B. (2000). A tantárgyakkal kapcsolatos attitűdök összefüggései. [Students' attitudes towards school subjects] *Magyar Pedagógia*, *100.* 343–366.

Csíkos, C. (2003). Matematikai szöveges feladatok megoldásának problémái 10–11 éves tanulók körében. [The difficulties of comprehending mathematical word problems among 10–11 year old students] *Magyar Pedagógia*, *103.* 35–55.

Davis-Dorsey, J., Ross, S. M., & Morrison, G. R. (1991). The role of rewording and context personalization in the solving of mathematical word problem solving. *Journal of Educational Psychology*, *83.* 61–68.

De Corte, E., & Verschaffel, L. (1981). Children's solution processes in elementary arithmetic problems: Analysis and improvement. *Journal of Educational Psychology*, *58*. 765–779.

De Lange, J. (1993). Between end and beginning: Mathematics education for 12–16 year olds: 1987–2002. *Educational Studies in Mathematics*, *25.* 137–160.

Depaepe, F., De Corte, E., & Verschaffel, L. (2009). Analysis of the realistic nature of word problems in upper elementary mathematics education. In L. Verschaffel, B. Greer, W. Van Dooren, & S. Mukhopadhyay (Eds.), *Words and worlds: Modelling verbal descriptions of situations* (pp. 245–264). Rotterdam: Sense Publications.

Doorman, L. M., & Gravemeijer, K. P. E. (2009). Emergent modeling: discrete graphs to support the understanding of change and velocity. *ZDM Mathematics Education*, *41.* 199–211.

Elbers, E., & de Haan, M. (2005). The construction of word meaning in a multicultural classroom. Mediational tools in peer collaboration during mathematics lessons. *European Journal of Psychology of Education*, *20.* 45–59.

Eriksson, G. (2008). Arithmetical thinking in children attending special schools for the intellectually disabled. *Journal of Mathematical Behavior*, *27.* 1–10.

Ernest, P. (1999). Forms of knowledge in mathematics and mathematics education: Philosophical and rhetorical perspectives. *Educational Studies in Mathematics*, *38.* 67–83.

Fitzpatrick, R., & Morrison, E. J. (1971). Performance and product evaluation. In R. L. Thorndike (Ed.), *Educational measurement*, (2nd ed.) (pp. 237–270). Washington, DC: American Council on Education.

Freudenthal, H. (1991). *Revisiting Mathematics Education. China Lectures*. Dordrecht: Kluwer.

Galbraith, P., & Stillman, G. (2001). Assumptions and context. Pursuing their role in modeling activity. In J.F. Matos, W. Blum, S.K. Houston, & S.P. Carreira (Eds.), *Modeling and mathematics education. ICTMA 9: Applications in science and technology* (pp. 300–310). Chichester, U.K.: Horwood.

Garcia, M., Sanchez, V., & Escudero, I. (2007). Learning through reflection in mathematics teacher education. *Educational Studies in Mathematics*, *64.* 1–17.

Gravemeijer, K. (1994). Educational development and developmental research in mathematics education. *Journal for Research in Mathematics Education*, *25.* 443–471.

Gravemeijer, K. (1997). Solving word problems: a case of modelling? *Learning and Instruction*, *7.* 389–397.
Gravemeijer, K., & Doorman, M. (1999). Context problems in realistic mathematics education: a calculus course as an example. *Educational Studies in Mathematics*, *39.* 111–129.
Gravemeijer, K., & Terwel, J. (2000). Hans Freudenthal: a mathematician on didactics and curriculum theory. *Journal of Curriculum Studies*, *32.* 777–796.
Greer, B. (1997). Modelling reality in mathematics classrooms: The case of word problems. *Learning and Instruction*, *7.* 293–307.
Guy, R. K. (1981). *Unsolved problems in number theory*. Springer-Verlag: New York – Heidelberg – Berlin.
Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education, 28.* 524–549.
Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K., Human, P., Murray, H., Olivier, A., & Wearne, D. (1996). Problem solving as a basis for reform in curriculum and instruction: The case of mathematics. *Educational Researcher*, *25.* 12–21.
High Level Group on Science Education (2007). *Science ducation now: A renewed pedagogy for the future of Europe*. European Commission: Brussels.
Hodgson, T., & Morandi, P. (1996). Exploration, explanation, formalization: A three-step approach to proof. *Primus, 6.* 49–57.
Hong, E. (1995). Mental models in word problem-solving: A comparison between American and Korean sixth-grade students. *Applied Cognitive Psychology*, *9.* 123–142.
Inoue, N. (2008). Minimalism as a guiding principle: Linking mathematical learning to everyday knowledge. *Mathematical Thinking and Learning*, *10.* 36–67.
Jitendra, A. K., Griffin, C. C., Deatline-Buchman, A., & Sczesniak, E. (2007). Mathematical word problem solving in third-grade classrooms. *The Journal of Educational Research*, *100.* 283–302.
Jitendra, A. K., Sczesniak, E., & Deatline-Buchman, A. (2005). An exploratory validation of curriculum-based mathematical word problem-solving tasks as indicators of mathematical proficiency for third graders. *School Psychology Review*, *34.* 358–371.
Jonassen, D. H. (2003). Designing research-based instruction for story problems. *Educational Psychology Review*, *15.* 267–296.
Keijzer, R., & Terwel, J. (2003). Learning for mathematical insight: a longitudinal comparative study of modelling. *Learning and Instruction*, *13.* 285–304.
Kintsch, W. (1985). Learning from text. *Cognition and Instruction*, *3.* 87–108.
Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review, 92.* 109–129
Klein, A. S., Beishuizen, M., & Treffers, A. (1998). The empty number line in Dutch second grades: realistic versus gradual program design. *Journal for Research in Mathematics Education*, *29.* 443–464.
Koninklijke Nederlandse Akademie van Wetenschappe [Royal Dutch Academia of Sciences] (2009). *Rekenonderwijs op de Basisschool. Analyse en Sleutels tot Verbetering*. [Mathematics education in the elementary school. Analysis and keys to improvement.] Koninklijke Nederlandse Akademie van Wetenschappen: Amsterdam.

Kramarski, B., Mevarech, Z. R., & Arami, M. (2002). The effects of metacognitive instruction on solving mathematical authentic tasks. *Educational Studies in Mathematics*, *49.* 225–250.

Kroesbergen, E. H., & van Luit, J. E. H. (2002). Teaching multiplication to low math performers: Guided versus structured instruction. *Instructional Science*, *30.* 361–378.

Lampert, M. (1986). Knowing, doing, and teaching multiplication. *Cognition and Instruction*, *3.* 305–342.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal*, *27.* 29–63.

Larsen, S., & Zandieh, M. (2008). Proofs and refutations in the undergraduate mathematics classroom. *Educational Studies in Mathematics*, *67.* 205–216.

Lave, J. (1992). Word problems: a microcosm of theories of learning. In P. Light, & G. Butterworth (Eds.), *Context and cognition. Ways of learning and knowing* (pp. 74–92). Hillsdale, NJ – Hove and London: Lawrence Erlbaum Associates.

Light, P., & Butterworth, G. (1992) (Ed.). *Context and cognition. Ways of learning and knowing*. Hillsdale, NJ – Hove and London: Lawrence Erlbaum Associates.

Linchevski, L., & Williams, J. (1999). Using intuition from everyday life in 'filling' the gap in children's extension of their number concept to include the negative numbers. *Educational Studies in Mathematics*, *39.* 131–147.

Mack, N. K. (1990). Learning fractions with understanding: Building on informal knowledge. *Journal for Research in Mathematics Education*, *21.* 16–32.

Maddy, P. (2008). How applied mathematics became pure. *The Review of Symbolic Logic, 1.* 16–41.

Morales, R. V., Shute, V. J., & Pellegrino, J. W. (1985). Developmental differences in understanding and solving simple mathematics word problems. *Cognition and Instruction*, *2.* 41–57.

Nemzeti alaptanterv (2007). [National Core Curriculum]

OECD (1999). *Measuring student knowledge and skills. A new framework for assessment*. Paris: OECD.

OECD (2003). *The PISA 2003 assessment framework – mathematics, reading, science, and problem solving. Knowledge and skills.* Paris: OECD.

OECD (2004). *First results from PISA 2003.* Paris: OECD.

OECD (2006). A*ssessing scientific, reading and mathematics literacy. A framework for PISA 2006*. Paris: OECD.

Palm, T. (2008). Impact of authenticity on sense making in word problem solving. *Educational Studies in Mathematics*, *67.* 37–58.

Palm, T. (2009). Theory of authentic task situations. In L. Verschaffel, B. Greer, W. Van Dooren, & S. Mukhopadhyay (Eds.), *Words and worlds: Modelling verbal descriptions of situations* (pp. 3–19.). Rotterdam: SensePublishers.

Pollak, H. O. (1969). How can we teach applications of mathematics? *Educational Studies in Mathematics*, *2.* 393–404.

Radatz, H. (1983). Untersuchungen zum Lösen einglekeideter Aufgaben. *Zeitschrift für Mathematik-Didaktik*, *3/83.* 205–217.

Reusser, K. (1988). Problem solving beyond the logic of things: contextual effects on understanding and solving word problems. *Instructional Science*, *17*. 309–338.

Reusser, K., & Stebler, R. (1997). Every word problem has a solution – the social rationality of mathematical modeling in schools. *Learning and Instruction*, *7*. 309–327.

Rickart, C. (1996). Structuralism and mathematical thinking. In R. J. Sternberg & T. Ben-Zeev (Eds.), The nature of mathematical thinking (pp. 285–300). Mahwah, NJ.: Lawrence Erlbaum Associates.

Riley, M. S., & Greeno, J. G. (1998). Developmental analysis of understanding language about quantities and of solving problems. *Cognition and Instruction*, *5*. 49–101.

Sáenz, C. (2009). The role of contextual, conceptual and procedural knowledge in activating mathematical competencies (PISA). *Educational Studies in Mathematics*, *71*. 123–143.

Saxe, G. B. (1988). The mathematics of child street vendors. *Child Development*, *59*. 1415–1425.

Säljö, R. (1991a). Culture and learning. *Learning and Instruction*, *1*. 179–185.

Säljö, R. (1991b). Learning and mediation: Fitting reality into a table. *Learning and Instruction*, *1*. 261–272.

Schoenfeld, A. H. (1988). When good teaching leads to bad results: The disasters of "well taught" mathematics courses. *Educational Psychologist*, *23*. 145–166.

Smolarski, D. C. (2002). Teaching mathematics in the seventeenth and twenty-first centuries. *Mathematics Magazine*, *75*. 256–262.

Sriraman, B., & Törner, G. (2008). Political union / mathematics education disunion. Building bridges in European didactic traditions. In L. D. English (Ed.), *Handbook of international research in mathematics education* (pp. 656–690). 2nd edition. New York and London: Routledge.

Stein, M. K., Remillard, J., & Smith, M. S. (2007). How curriculum influences student learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 319–369). Charlotte, NC: Information Age.

Sternberg, R. J., & Frensch, P. A. (1992). On being an expert: A cost-benefit analysis. In R. R. Hoffman (Ed.), *The psychology of expertise: Cognitive research and empirical AI* (pp. 191–204). New York, NY: Springer.

Streefland, L., & van den Heuvel-Panhuizen, M. (1999). Uncertainty, a metaphor for mathematics education? *Journal of Mathematical Behavior*, *17*. 393–397.

Szendrei, J. (2007). When the going gets tough, the tough gets going problem solving in Hungary, 1970–2007: research and theory, practice and politics. *ZDM Mathematics Education*, *39*. 443–458.

Treffers, A. (1993). Wiscobas and Freudenthal realistic mathematics education. *Educational Studies in Mathematics*, *25*(1–2). 89–108.

van den Heuvel-Panhuizen, M. (1996). *Assessment and realistic mathematics education*. Utrecht, The Netherlands: CD-β Press.

van den Heuvel-Panhuizen, M. (2000). *Mathematics education in the Netherlands: A guided tour*. Freudenthal Institute Cd-rom for ICME9. Utrecht: Utrecht University.

van den Huivel-Panhuizen, M. (2001a). Realistic Mathematics Education as work in progress. In F. L. Lin (Ed.), *Common Sense in Mathematics Education. Proceedings of 2001 The Netherlands and Taiwan Conference on Mathematics Education* (pp. 1–43)., Taipei, Taiwan.

van den Huivel-Panhuizen, M. (2001b). The role of contexts in assessment problems in mathematics. *For the Learning of Mathematics*, *25.* 2–9.

van den Heuvel-Panhuizen, M. (2003). The didactical use of models in realistic mathematics education: An example from a longitudinal trajectory on percentage. *Educational Studies in mathematics*, *54.* 9–35.

van Garderen, D. (2004). Reciprocal teaching as a comprehension strategy for understanding mathematical word problems. *Reading & Writing Quarterly*, *20.* 225–229.

van Garderen, D. (2007). Teaching students with LD to use diagrams to solve mathematical word problems. J*ournal of Learning Disabilities*, *40.* 540–553.

Verschaffel, L. (2002). Taking the modeling perspective seriously at the elementary school level: promises and pitfalls (Plenary lecture). In A. Cockburn & E. Nardi (Eds.), *Proceedings of the 26th Annual Conference of the International Group for the Psychology of Mathematics Education.* (Vol. 1, pp. 64–82). School of Education and Professional Development, University of East Anglia, UK.

Verschaffel, L., De Corte, E., & Borghart, I. (1997). Pre-service teachers' conceptions and beliefs about the role of real-world knowledge in mathematical modeling of school word problems. *Learning and Instruction*, *7.* 339–360.

Verschaffel, L., De Corte, E. & Lasure, S. (1994). Realistic considerations in mathematical modeling of school arithmetic word problems. *Learning and Instruction*, *7.* 339–359.

Verschaffel, L., Greer, B., & De Corte, E. (2000). *Making sense of word problems*. Lisse, The Netherlands: Swets & Zeitlinger.

Wubbels, T., Korthagen, F., & Broekman, H. (1997). Preparing teachers for realistic mathematics education. *Educational Studies in Mathematics*, *32*. 1–28.

Wyndhamn, J. & Säljö, R. (1997). Word problems and mathematical reasoning – a study of children's mastery of reference and meaning in textual realities. *Learning and Instruction*, *7*. 361–382.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27.* 458–477.